AD-A069 571 STANFORD UNIV CALIF DEPT OF STATISTICS F/G 9/4

EFFICIENT ESTIMATION OF A MODEL WITH AN AUTOREGRESSIVE SIGNAL W--ETC(U)

MAR 79 Y HOSOYA N00014-75-C-0442

UNCLASSIFIED TR-37 NL

1 OF 1

AD A069571

END

DATE FILMED

7 -79

DDC

(12)

LEVEL II

AD A069571

EFFICIENT ESTIMATION OF A MODEL WITH AN AUTOREGRESSIVE SIGNAL WITH WHITE NOISE

BY

YUZO HOSOYA

TECHNICAL REPORT NO. 37
MARCH 1979

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 37	2. GOVT ACCESSION NO. 14 TR-37	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 EFFICIENT ESTIMATION OF A MODEL WITH AN AUTOREGRESSIVE SIGNAL WITH WHITE NOISE.		5. TYPE OF REPORT & PERIOD COVERED 9 Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Y. HOSOYA		8. CONTRACT OR GRANT NUMBER(s) 15 N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California 332 580		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE 11 MARCH 1979
		13. NUMBER OF PAGES 34
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 10 Yuzo / Hosoya		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

12 42 p.

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

autoregressive signal plus white noise, approximate maximum-likelihood estimate, Whittle-Walker model, asymptotic distribution, Newton-Raphson iteration

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

332 580

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT.

This paper considers the estimation of parameters in the model of $X_t = S_t + \varepsilon_t$ where the S_t are generated by a stationary autoregressive model $\sum_{i=0}^{p} \alpha_i S_{t-i} = \eta_t$ and the η_t and the ε_t are i.i.d. random variables. In case the η_t and the ε_t are Gaussian, Hosoya (Yale Ph.D. thesis, 1974), Pagano (Ann. Stat., 1974) and Dunsmuir (Ann. Stat., 1979), respectively, constructed efficient estimates and gave their asymptotic distribution. This paper gives the asymptotic distribution of an approximate maximum-likelihood estimate using only a condition on the fourth-order moments of ε_t and η_t and without the assumption of normality. This paper also contains a theorem which shows that under general conditions an estimate given by the second-step in the Newton-Raphson iteration with a consistent estimate as an initial value is second-order efficient in view of C. R. Rao's definition (Rao, J.R.S.S.B., 1962).

epsilon

sub t

eta sub t

epsilon sub t

sum from i=0 to p of (alpha sub i)(S sub t-i)= eta sub t

Accession For	
NTIS GRA&I	☒
DDC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist. A	Avail and/or special

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

EFFICIENT ESTIMATION OF A MODEL WITH AN AUTOREGRESSIVE SIGNAL WITH WHITE NOISE

by

YUZO HOSOYA

Technical Report No. 37

March 1979

PREPARED UNDER CONTRACT N00014-75-C-0442

(NR-042-034)

OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Efficient Estimation of a Model with an Autoregressive Signal with White Noise

by

Yuzo Hosoya

Tohoku University
Japan

Abstract

This paper considers the estimation of parameters in the model of $X_t = S_t + \varepsilon_t$ where the S_t are generated by a stationary autoregressive model $\sum_{i=0}^{p} \alpha_i S_{t-i} = \eta_t$ and the η_t and the ε_t are i.i.d. random variables. In case the η_t and the ε_t are Gaussian, Hosoya (Yale Ph.D. thesis, 1974),Pagano (Ann. Stat., 1974) and Dunsmuir (Ann. Stat., 1979), respectively, constructed efficient estimates and gave their asymptotic distribution. This paper gives the asymptotic distribution of an approximate maximum-likelihood estimate using only a condition on the fourth-order moments of ε_t and η_t and without the assumption of normality. This paper also contains a theorem which shows that under general conditions an estimate given by the second-step in the Newton-Raphson iteration with a consistent estimate as an initial value is second-order efficient in view of C. R. Rao's definition (Rao, J.R.S.S.B., 1962).

Key words: autoregressive signal plus white noise, approximate maximum-likelihood estimate, Whittle-Walker model, asymptotic distribution, Newton-Raphson iteration.

Efficient Estimation of a Model with
an Autoregressive Signal with White Noise

by

Yuzo Hosoya
Tohoku University
Japan

0. Introduction.

Suppose that a message X_t has been received, but that, due to noise in the channel of communication, the original signal s_t cannot be reconstructed directly from the observation X_t $(=s_t+\epsilon_t)$. The techniques of so-called signal detection (or signal extraction) have been developed for the purpose of inferring the signal sent as an important field of communication theory. [See Whalen (1971).] This same problem has also been called, by some econometricians, the problem of "unobservable variables". Namely, they maintain that the actually observed quantities do not necessarily coincide with the corresponding variables in a theoretical framework; thus certain noise-elimination techniques need to be applied to observations when a theoretical model is fitted. [See, for example, Grether and Nerlove (1970).]

In a probabilistic framework, this signal-extraction problem has a direct connection with prediction theory in stationary stochastic processes, and a similar technique to that of construction of the optimal linear filter in prediction can be applied. In particular, if the spectral densities of the s_t and the ϵ_t are rational, the optimal (in the sense of minimum

---

A part of this work was done while I was visiting the Department of Statistics at Stanford University during the fall quarter of 1978. I would like to express my sincerest thanks to Professor T. W. Anderson for his reading and improving the paper, and also to Messrs. S. Sugihara and F. Ahrabi for their pertinent comments.

mean-square error) estimate of s_t can be obtained from a fairly simple recursive formula. [See Whittle (1963).] Prediction theory, however, assumes complete knowledge of the spectral structure of both signal and noise. However, in practical situations, this is not usually the case. Rather, in most cases, what is required is to recover the information concerning the structure of the signal and noise. For this purpose, there seems to be two statistical treatments. One is to assume s_t to be a certain (not necessarily linear) deterministic function of a time parameter or of other parameters and to apply least-squares or other pertinent methods (Walker (1969) and Hannan (1971)). Another approach considers the model of a nondeterministic stationary signal. This paper explores the latter approach. The model which will be investigated below is the following: Assume that a signal s_t is observed superimposed by white noise ε_t , that is,

(1) $$X_t = s_t + \varepsilon_t ,$$

and assume further that the signal is generated by an autoregressive process

(2) $$\sum_{i=0}^{p} \alpha_i s_{t-i} = \eta_t , \quad t = \ldots, -1, 0, 1, \ldots,$$

with $\alpha_0 \equiv 1$, where the ε_t and the η_t are respectively i.i.d., and mutually independent. Henceforth write α the vector whose element is α_j .

The interest in investigating specifically this type of model is that, as will be seen, this is a special case of rational spectra. [The

spectrum of the present model is formally a rational function of $e^{i\omega}$, but the parameters of the denominator and of the numerator are functionally related to each other. The usual statistical estimation procedures of rational spectra do not seem to apply to the present model effectively. Though the author attempted the application of Hannan's method of estimation of rational spectra (see Hannan (1970)) to this model, it was unsatisfactory. It seems that only when the variance ratio of ε_t and η_t is known, can Hannan's method (after certain modifications) be applied.]

In his paper Parzen (1967) suggests using the Yule-Walker equations or the instrumental variable method to estimate the parameters in the model expressed by (1) and (2). The method is consistent, but not efficient. The Yule-Walker equations can be derived as follows. Noting that, in view of (1) and (2), $\sum_{i=0}^{p} \alpha_i E(X_{t-i} X_{t-p-\ell}) = 0$ for $\ell = 1, 2, \ldots, p$, an estimate of the α_i can be obtained by solving those equations after replacing $E(X_{t-i} X_{t-p-\ell})$ by the sample covariance $\sum_{t=p+\ell+1}^{N} X_{t-i} X_{t-p-\ell}/N-p-\ell$. Walker (1960) observed that in the case $p = 1$ the efficiency of this estimate is near unity only for a small α_1 or for a high signal to noise proportion.

This paper considers the model given by (1) and (2), and establishes the asymptotic properties of an approximate maximum-likelihood estimate and an efficient estimate is also constructed. An estimate is called efficient below when its asymptotic distribution is normal with asymptotic covariance matrix equal to the limit of the inverse of the average Fisher information matrix when the process is Gaussian. Another approach for constructing an efficient estimate of the α's was proposed by Pagano (1974),

though his estimate is different from the one given here in that his method does not use the likelihood function and also his consistent estimate for the starting value of iteration is different from the one proposed below. The author of the present paper has shown an optimality of the use of likelihood function to obtain an efficient estimate of parameters in time-series models in his paper (1977).

The program proceeds as follows: In the next section, an approximate likelihood function is given, and the asymptotic properties of the approximate maximum-likelihood estimate are derived as a corollary of a more general result concerning the approximate maximum-likelihood estimation of a linear process plus white noise; the result is given in Appendix 2 since it has an independent interest. [For the simplicity of terminology, the value of the parameter maximizing an approximate likelihood function will be called below the maximum-likelihood estimate. This will cause no confusion.] Section 2 concerns the construction of an efficient estimate of α , σ_ε^2 and σ_η^2 where σ_ε^2 and σ_η^2 are respectively the variance of ε_t and η_t . A method that yields consistent estimates of σ_ε^2 and σ_η^2 is shown and an efficient estimate of α will be given by the Newton-Raphson iteration. Appendix 1 to this paper establishes that under general conditions the second Newton-Raphson iteration gives an estimate which is equivalent to the maximum-likelihood estimate to the probability order $1/N$, where N is the sample size, whereas efficient estimates in general are equivalent to the maximum-likelihood estimate to the probability order $1/\sqrt{N}$.

Finally, since this paper exclusively considers the case where a signal is generated by an autoregressive scheme, it may be pertinent to

offer a comment on the moving-average type signal. Consider, for example, the simplest moving-average scheme $s_t = \eta_t + \alpha\eta_{t-1}$ and suppose the X_t $(= s_t + \varepsilon_t)$ are observed. Assuming the same conditions on ε_t and η_t as previously given, the spectral density of X_t is given by $f(\omega|\alpha, \sigma_\eta^2, \sigma_\varepsilon^2) = \frac{1}{2\pi}\{|1 + \alpha e^{i\omega}|^2 \sigma_\eta^2 + \sigma_\varepsilon^2\}$, $-\pi \leq \omega < \pi$. If σ_ε^2 and σ_η^2 are unknown, the values of α , σ_ε^2 and σ_η^2 that give the same spectral density f as a function of ω are not unique. Suppose $|1 + \alpha e^{i\omega}|^2 \sigma_\eta^2 + \sigma_\varepsilon^2 = |1 + \alpha^* e^{i\omega}|^2 \sigma_\eta^{*2} + \sigma_\varepsilon^{*2}$, from which it follows that

$$(1 + \alpha^2)\, \sigma_\eta^2 + \sigma_\varepsilon^2 = (1 + \alpha^{*2})\, \sigma_\eta^{*2} + \sigma_\varepsilon^{*2} ,$$

$$\alpha\sigma_\eta^2 = \alpha^* \alpha_\eta^{*2} .$$

Given α^*, σ_ε^* and σ_η^{*2} , it is easy to see that the solution of the equations above is indeterminate, even if there is the restriction $|\alpha| < 1$.

1. The Likelihood Function.

An approximate likelihood function for α, σ_ε^2 and σ_η^2 is derived here under the assumption that ε_t and η_t are Gaussian. First of all the spectral density of X_t generated by (1) and (2) can be written as

$$(3) \qquad f(\omega|\alpha, \sigma_\varepsilon^2, \sigma_\eta^2) = \frac{1}{2\pi}\left[\frac{\sigma_\eta^2}{\left|\sum_{j=0}^{p} \alpha_j e^{ij\omega}\right|^2} + \sigma_\varepsilon^2\right] , \; -\pi \leq \omega < \pi ,$$

where $\alpha_0 = 1$. Write (3) as $\left\{\sigma_\eta^2 + \sigma_\varepsilon^2 \left| \Sigma_j \alpha_j e^{ij\omega}\right|^2\right\} \Big/ 2\pi \left| \Sigma_j \alpha_j e^{ij\omega}\right|^2$.

Then the numerator can be factorized into $\sigma^2\left|\sum_{j=0}^{p} \beta_j e^{ij\omega}\right|^2$ for some σ^2 and β $(\beta_0 = 1)$. This immediately follows from the Fejér-Riesz theorem [for example, Akhiezer (1956), p. 152] which says that if

$$(4) \qquad g(\omega) = \sum_{k=-p}^{p} \alpha_k e^{ik\omega}, \ -\pi \leq \omega < \pi,$$

and $g(\omega)$ is real and nonnegative, then there exists an $h(\omega)$ such that $g(\omega) = |h(\omega)|^2$ and $h(\omega) = \sum_{k=0}^{p} \beta_k e^{ik\omega}$. The numerator of (3) is, as is obvious from its expansion, of the form (4) and nonnegative; therefore

$$(5) \qquad f(\omega|\alpha, \sigma_\varepsilon^2, \sigma_\eta^2) = \frac{1}{2\pi} \frac{\sigma^2\left|\Sigma_k \beta_k e^{ik\omega}\right|^2}{\left|\Sigma_j \alpha_j e^{ij\omega}\right|^2},$$

where σ^2 and β are functions of $\alpha, \sigma_\varepsilon^2$ and σ_η^2.

Now in view of (5), X_t may be interpreted as generated by a linear process $X_t = \sum_{k=0}^{\infty} \nu_k e_{t-k}$, where the ν_k's are obtained from the equations $\sum_{k=0}^{\infty} \nu_k e^{i\omega k} = \Sigma_\ell \beta_\ell e^{i\ell\omega} / \Sigma_j \alpha_j e^{ij\omega}$, and where the e_t are independent random variables such that

$$(6) \qquad \operatorname{Var}(e_t) = 2\pi \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega) d\omega$$

$$= 2\pi \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\log\left(\frac{\sigma^2}{2\pi}\right) + \log \frac{\left|\Sigma_k \beta_k e^{ik\omega}\right|^2}{\left|\Sigma_j \alpha_j e^{ij\omega}\right|^2} \right] d\omega$$

$$= \sigma^2 \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log \frac{\left|\Sigma_k \beta_k e^{ik\omega}\right|^2}{\left|\Sigma_j \alpha_j e^{ij\omega}\right|^2} d\omega\right\}$$

$$= \sigma^2 .$$

Also in view of (3), (6) may be written as

$$\sigma^2 = 2\pi \exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left\{\frac{\sigma_\eta^2}{\left|\Sigma_j \alpha_j e^{ij\omega}\right|^2} + \sigma_\varepsilon^2\right\} d\omega\right] \tag{7}$$

The equation (6) is due to Szego's theorem [c.f. Hoffman (1962)].

Assuming the normality of ε_t and η_t, after $X_1, X_2, \ldots, X_N$ are observed, the log-likelihood function is

$$\log L_N(\alpha, \sigma_\varepsilon^2, \sigma_\eta^2) = -\frac{1}{2}\log|V_N| - \frac{N}{2}\log 2\pi\sigma^2 - Q_N(X, \alpha, \sigma_\varepsilon^2, \sigma_\eta^2) , \tag{8}$$

where $\sigma^2 V_N$ is the variance-covariance matrix of $X_1, X_2, \ldots, X_N$, and

$$Q_N(X, \alpha, \sigma_\varepsilon^2, \sigma_\eta^2) = X' V_N^{-1} X/(2\sigma^2) .$$

The log-likelihood function (8) can be simplified by making use of the following results due to Whittle (1952, 1962).

A) For a linear process $X_t = \sum_{i=0}^{\infty} \eta_i \varepsilon_{t-i}$, if $\Sigma_k \eta_k Z^k$ is analytic and non-zero on $\{Z:|Z| < 1 + \delta\}$ for some $\delta > 0$, then $\log|V_N| \to 0$ as $N \to \infty$. The present model satisfies this condition. From the formation of β, it is obvious that $\sigma^2|\Sigma\beta_k Z^k|^2 = \sigma_\eta^2 + \sigma_\varepsilon^2 |\Sigma_j \alpha_j Z^j|^2 \neq 0$. Also all zeroes of $\Sigma_k \alpha_k Z^k = 0$ are outside the unit circle. Therefore $1/(\Sigma_k \alpha_k Z^k)$

is analytic and $|\Sigma\alpha_j Z^j| \neq 0$. Thus $\Sigma\eta_k Z^k = \Sigma\beta_j Z^j \big/ \Sigma\alpha_\ell Z^\ell$ is analytic and nonzero on $\{Z: |Z| < 1 + \delta\}$ for some δ .

B) For large N, Q_N can be approximated by

$$(9) \quad U_N(X, \alpha, \sigma_\varepsilon^2, \sigma_\eta^2) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \left|\sum_{n=1}^{N} X_n e^{in\omega}\right|^2 \Big/ f(\omega|\alpha, \sigma_\varepsilon^2, \sigma_\eta^2)\, d\omega$$

$$= N \sum_{k=-(N-1)}^{N-1} \delta_k C_k ,$$

where σ_k is the k-th Fourier coefficient of $1?f(\omega|\alpha, \sigma_\varepsilon^2, \sigma_\eta^2)$ and C_k is the sampling autocovariance of k-th order.

Now by A) and B), the log-likelihood function (8) can be approximated for large N by

$$(10) \quad \log L_N^*(\theta) = -\frac{N}{2} \log 2\pi\sigma^2(\theta) - \frac{1}{N} U_N(X,\theta)$$

$$= -\frac{N}{2} \log (2\pi)^2 - \frac{N}{2}\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega|\theta) d\omega$$

$$- \frac{1}{2(2\pi)^2} \int_{-\pi}^{\pi} \frac{|\Sigma_n X_n e^{i\omega n}|^2}{f(\omega|\theta)} d\omega ,$$

where $\theta_i = \alpha_i$ for $i = 1, 2, \ldots, p$ and $\theta_{p+1} = \sigma_\varepsilon^2$, $\theta_{p+2} = \sigma_\eta^2$, and thus $f(\omega|\theta) = \frac{1}{2\pi} \left(\theta_{p+2} \Big/ \left|\sum_{i=0}^{p} \theta_j e^{ij\omega}\right|^2 + \theta_{p+1} \right)$. The general treatment of the asymptotic properties of the least-squares estimate obtained from

maximizing the approximate likelihood function of the form (10) is furnished in Appendix 2, and the following Theorem 1 is a straightforward consequence of that result. That appendix deals with the asymptotic properties of the least-squares estimate of parameters of a general linear process which is superimposed by a white noise and derives them by means of an extension of the Whittle-Walker theorem. [See Whittle (1952) and Walker (1964).]

For the model represented by (1) and (2), assume the following:

A-1) The ε_t and η_t are strictly stationary processes with finite fourth cumulants which are denoted as $K_4(\varepsilon)$ and $K_4(\eta)$ respectively.

A-2) Let α^0, $\sigma^2_{\varepsilon 0}$, and $\sigma^2_{\eta 0}$ be the true values of α, σ^2_ε, σ^2_η respectively. Then $\alpha^0 \in A$ with A a compact subset of R^p such that, for any $\alpha \in A$, all zeroes of $\sum_{i=0}^{p} \alpha_i Z^i$ are outside of the unit circle. $\sigma^2_{\varepsilon 0}$ and $\sigma^2_{\eta 0}$ are respectively in a compact subset of R^+.

Then

<u>Theorem 1</u>:

Let $\hat{\alpha}$, $\hat{\sigma}^2_\varepsilon$ and $\hat{\sigma}^2_\eta$ be the least-squares estimates derived from the function (10). Let $h(\omega|\theta) = 1/f(\omega|\theta)$. Then $\sqrt{N}(\hat{\alpha} - \alpha_0)$, $\sqrt{N}(\hat{\sigma}^2_\varepsilon - \sigma^2_{\varepsilon 0})$ and $\sqrt{N}(\hat{\sigma}^2_\eta - \sigma^2_{\eta 0})$ are asymptotically jointly normally distributed with mean vector 0, and with covariance matrix $4\pi W_0^{-1} + K_4(\varepsilon) W_0^{-1} U_0 W_0^{-1} + K_4(\eta) W_0^{-1} V_0 W_0^{-1}$, where W_0, U_0 and V_0 are $(p+2)\times(p+2)$ matrices with the representative terms

$$\int_{-\pi}^{\pi} \frac{h^{(i)}(\omega|\theta^0)}{h(\omega|\theta^0)} \frac{h^{(j)}(\omega|\theta^0)}{h(\omega|\theta^0)} d\omega \ , \quad \int_{-\pi}^{\pi} h^{(i)}(\omega|\theta^0) \frac{\theta^0_{p+1}}{2\pi g(\omega|\theta^0)} d\omega$$

$$\int_{-\pi}^{\pi} h^{(j)}(\omega|\theta^0) \frac{\theta^0_{p+1}}{2\pi(\omega|\theta^0)} d\omega \ , \ \frac{\theta^0_{p+2}}{2\pi} \int_{-\pi}^{\pi} h^{(i)}(\omega|\theta^0) d\omega$$

$\frac{\theta^0_{p+2}}{2\pi} \int_{-\pi}^{\pi} h^{(j)}(\omega|\theta^0)\, d\omega$, respectively and where $\theta^0_i = \alpha^0_i$, $i = 1,2,\ldots,p$, $\theta^0_{p+1} = \sigma^2_{\varepsilon 0}$ and $\theta^0_{p+2} = \sigma^2_{\eta 0}$.

2. Efficient estimates of α , σ^2_ε and σ^2_η .

To obtain the maximum-likelihood estimate, the function to be maximized is, in view of (10),

$$-\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(\omega|\theta)d\omega - \frac{1}{(2\pi)^2 N}\int_{-\pi}^{\pi} \frac{\left|\Sigma_n X_n e^{i\omega n}\right|^2}{f(\omega|\theta)} d\omega$$

which, for practical purpose, can be approximated by

$$- \sum_{i=0}^{N-1} \log f(\omega_i|\theta) - \sum_{i=0}^{N-1} \frac{I(\omega_j)}{f(\omega_j|\theta)} \ , \tag{11}$$

where $\omega_j = 2\pi j/N$, $j = 0, 1, \ldots, N-1$, and $I(\omega_j) = \frac{1}{2\pi N}\left|\Sigma_n X_n e^{i\omega_j n}\right|^2$.

Let the quantity (11) be denoted by $A(\theta,X)$. The first derivative of $A(\theta,X)$ is nonlinear with respect to θ , so that a certain approximation

is required for the solution of $\partial A(\theta,X)/\partial\theta = 0$. It can be shown that the Newton-Raphson iteration procedure generally produces an estimate as efficient as the maximum-likelihood estimate if the iteration starts with a consistent estimate of θ ; Theorem 3 of Appendix 1 proves that the first-step iteration produces an estimate θ^2 such that $\sqrt{N}(\theta^2-\theta)$ converges to 0 in probability and moreover that $N(\theta^3-\theta)$ tends to 0 in probability for the estimate θ^3 obtained by the second-step of the iteration. For that theorem to apply, two points must be checked. The one is whether the starting value of θ is consistent, and the other is whether the present model satisfies the conditions of Theorem 3. Concerning the first point, it has already been shown that the solution of the Yule-Walker equations is a consistent estimate of α . The starting value for σ_ε^2 and σ_η^2 can be constructed as follows: Let $g(\omega|\alpha) = \left|\sum_{k=0}^{p} \alpha_k e^{i\omega k}\right|^2$, and let $\tilde{\alpha}$ be the solution of the Yule-Walker equations. Taking relation (3) into consideration, regress $2\pi I(\omega_j)$ on $1/g(\omega_j|\tilde{\alpha})$, $j = 0, 1, \ldots, N-1$. Then estimates of σ_ε^2 and σ_η^2 can be obtained as the regression coefficients. Namely, calculate

(12) $$\tilde{\sigma}_\varepsilon^2 = \frac{\frac{1}{N}\sum_j \frac{2\pi I(\omega_j)}{g(\omega_j|\tilde{\alpha})} - \frac{1}{N}\sum_j 2\pi I(\omega_j)\,\frac{1}{N}\sum_j \frac{1}{g(\omega_j|\tilde{\alpha})}}{\frac{1}{N}\sum_j \frac{1}{g(\omega_j|\tilde{\alpha})} - \left[\frac{1}{N}\sum \frac{1}{g(\omega_j|\tilde{\alpha})}\right]^2} ,$$

(13) $$\tilde{\sigma}_\eta^2 = \frac{1}{N}\sum_j 2\pi I(\omega_j) - \tilde{\sigma}_\varepsilon^2 \frac{1}{N}\sum_j 1/g(\omega_j|\tilde{\alpha}) .$$

Theorem 2:

The $\tilde{\sigma}_\varepsilon^2$ and $\tilde{\sigma}_\eta^2$ are consistent estimates of σ_ε^2 and σ_η^2 respectively.

Proof:

$$\operatorname*{plim}_{N\to\infty} \tilde{\sigma}_\varepsilon^2 = \frac{\frac{1}{2\pi}\int \frac{2\pi f(\omega|\alpha, \sigma_\varepsilon^2, \sigma_\eta^2)}{g(\omega|\alpha)} d\omega - \frac{1}{2\pi}\int 2\pi f(\omega|\alpha, \sigma_\varepsilon^2, \sigma_\eta^2) d\omega \, \frac{1}{2\pi}\int \frac{1}{g(\omega|\alpha)} d\omega}{\frac{1}{2\pi}\int \frac{1}{g(\omega|\alpha)^2} d\omega - \left\{ \frac{1}{2\pi}\int \frac{1}{g(\omega|\alpha)} d\omega \right\}^2}$$

where, using (3), the numerator above is equal to

$$\frac{\sigma_\varepsilon^2}{2\pi}\int \frac{d\omega}{g(\omega|\alpha)^2} + \frac{\sigma_\eta^2}{2\pi}\int \frac{d\omega}{g(\omega|\alpha)} - \frac{1}{2\pi}\left\{ \int \frac{\sigma_\varepsilon^2 \, d\omega}{g(\omega|\alpha)} + 2\pi\sigma_\eta^2 \right\} \frac{1}{2\pi}\int \frac{d\omega}{g(\omega|\alpha)}$$

$$= \frac{\sigma_\varepsilon^2}{2\pi} \left\{ \frac{d\omega}{g(\omega|\alpha)^2} - \left(\frac{1}{2\pi} \quad \frac{d\omega}{g(\omega|\alpha)} \right)^2 \right\} .$$

Thus $\operatorname*{plim}_{N} \tilde{\sigma}_\varepsilon^2 = \sigma_\varepsilon^2$. In the same way, $\operatorname*{plim}_{N\to\infty} \tilde{\sigma}_\eta^2 = \sigma_\eta^2$. ///

In order to see that the Newton-Raphson method with $\tilde{\alpha}$, $\tilde{\sigma}_\varepsilon^2$ and $\tilde{\sigma}_\eta^2$ as its starting values provides an efficient estimate, it is sufficient to check that the approximate likelihood function (11) with conditions A-1 and A-2 satisfies conditions C-1 and C-2 of Theorem 3. Condition A-2 implies that $f(\omega|\theta)$, $\partial^2 f(\omega|\theta)/\partial\theta_i\partial\theta_j$ and $\partial^3 f(\omega|\theta)/\partial\theta_i\partial\theta_j\partial\theta_h$, $i, j, h = 1, 2, \ldots, p+2$, are uniformly continuous with respect to $\omega \in (-\pi,\pi)$ and θ in the parameter space. Thus convergence of $\frac{\partial^2 A(\theta,X)}{N\partial\theta_i\partial\theta_j}$ and $\frac{\partial^3 A(\theta,X)}{N\partial\theta_i\partial\theta_j\partial\theta_h}$ is straightforward. Accordingly conditions C-1 and C-2 are satisfied.

To summarize the preceding argument, efficient estimates of α, σ_ε^2 and σ_η^2 can be constructed as follows:

1) Solve $\sum_{i=0}^{p} \alpha_i \left(\sum_{t=p+\ell+1}^{N} X_{t-i} X_{t-p-\ell} \Big/ N-p-\ell \right) = 0$. $\ell = 1, 2, \ldots, p$

   for α . Let the solution be $\tilde{\alpha}$.

2) Calculate $\tilde{\sigma}_\varepsilon^2$ and $\tilde{\sigma}_\eta^2$ by (12) and (13).

3) Let $\theta_i^{(1)} = \tilde{\alpha}_i$, $i = 1, 2, \ldots, p$, $\theta_{p+1}^{(1)} = \tilde{\sigma}_\varepsilon^2$ and $\theta_{p+2}^{(1)} = \tilde{\sigma}_\eta^2$ and

   apply the following iteration formula:

$$\theta^{(n)} = \theta^{(n-1)} - \frac{\partial^2 A(\theta^{(n-1)},X)}{\partial\theta\,\partial\theta'}^{-1} \frac{\partial A(\theta^{(n-1)},X)}{\partial\theta}, \quad n = 2, 3, \ldots,$$

where ($A(\theta,X)$ and $g(\omega|\alpha)$ below are abbreviated as A and g)

$$\frac{\partial A}{\partial\theta_\ell} = \frac{\partial A}{\partial\alpha_\ell}$$

$$= -\sum_j \frac{2\sigma_\eta^2(\sum_k \alpha_k \cos(k-\ell)\omega_j)}{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)g} - \sum_j \frac{2I(\omega_j)\ {}^2(\sum_k \alpha_k \cos(k-\ell)\omega_j)}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^2}, \quad \ell = 1, 2, \ldots, p,$$

$$\frac{\theta A}{\partial\theta_{p+1}} = \frac{\partial A}{\partial\sigma_\varepsilon^2} = -\sum_j \frac{g}{\sigma_\eta^2 + \sigma_\varepsilon^2 g} + \sum_j \frac{I(\omega_j)g^2}{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)^2},$$

$$\frac{\partial A}{\partial\theta_{p+2}} = \frac{\partial A}{\partial\sigma_\eta^2} = -\sum_j \frac{1}{\sigma_\eta^2 + \sigma_\varepsilon^2 g} + \sum_j \frac{I(\omega_j)g}{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)^2},$$

$$\frac{\partial^2 A}{\partial\theta_\ell \partial\theta_m} = \frac{\partial^2 A}{\partial\alpha_\ell \partial\alpha_m}$$

$$= -\sum_j \frac{2\sigma_\eta^2 \cos\{(m-\ell)\omega_j\}}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^2} + \sum_j \frac{4\sigma_\eta^2(\sigma_\eta^2 + 2\sigma_\varepsilon^2 g)(\sum \alpha_k \cos(k-\ell)\omega_j)(\sum_k \alpha_k \cos(k-m)\omega_j)}{\{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)g\}^2}$$

$$-\sum_j \frac{2I(\omega_j)\sigma_\eta^2 \cos(m-\ell)\omega_j}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^2} + \sum_j \frac{4I(\omega_j)\sigma_\eta^2(\sum_k \alpha_k \cos(k-\ell)\omega_j)(\sum_k \alpha_k \cos(k-m)\omega_j)}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^3}$$

$$\frac{\partial^2 A}{\partial\theta_\ell \partial\theta_{p+1}} = \frac{\partial^2 A}{\partial\theta_\ell \partial\sigma_\varepsilon^2}$$

$$= \sum_j \frac{2\sigma_\eta^2(\sum_k \alpha_k \cos(k-\ell)\omega_j)}{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)^2} + \sum_j \frac{4I(\omega_j)\sigma_\eta^2 g(\sum_k \alpha_k \cos(k-\ell)\omega_k)}{\;}$$

$$\frac{\partial^2 A}{\partial\theta_{p+1}^2} = \frac{\partial^2 A}{\partial(\sigma_\varepsilon^2)^2}$$

$$= \sum_j \frac{g^2}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^2} - \sum_j \frac{2I(\omega_j)g^3(\sigma_\varepsilon^2 + \sigma_\eta^2 g)}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^4}$$

$$\frac{\partial^2 A}{\partial\theta_{p+2}^2} = \frac{\partial^2 A}{\partial(\sigma_\eta^2)^2}$$

$$= \sum_j \frac{1}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^2} - \sum_j \frac{2I(\omega_j)g}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^3}$$

$$\frac{\partial^2 A}{\partial\theta_\ell \partial\theta_{p+2}} = \frac{\partial^2 A}{\partial\alpha_\ell \partial\sigma_\eta^2}$$

$$= - \sum_j \frac{2\sum \alpha_k \cos(k-\ell)\omega_j}{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)^2} + \sum_j \frac{2\sigma_\eta^2(\sum_k \alpha_k \cos(k-\ell)\omega_j)}{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)^2}$$

$$- \sum_j \frac{2I(\omega_j)\sum_k \alpha_k \cos(k-\ell)\omega_j}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^2} + \sum_j \frac{4I(\omega_j) g\sigma_\eta^2(\sum_k \alpha_k \cos(k-\ell)\omega_m)}{(\sigma_\varepsilon^2 + \sigma_\eta^2 g)^3}$$

$$\frac{\partial^2 A}{\partial\theta_{p+1} \partial\theta_{p+2}} = \frac{\partial^2 A}{\partial\sigma_\varepsilon^2 \partial\sigma_\eta^2}$$

$$= \sum_j \frac{g}{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)^2} - \sum_j \frac{2I(\omega_j) g^2}{(\sigma_\eta^2 + \sigma_\varepsilon^2 g)^3}$$

3. Numerical Examples.

Here are given two examples of the method for the simulated data generated from the model

$$X_t = 0.6X_{t-1} + \eta_t \; ,$$

$$(14) \qquad Y_t = X_t + \varepsilon_t \; ,$$

where ε_t and η_t are respectively generated by a normal random-number generator with mean 0 and standard deviation 1 . The numerical results

in Table 1 exhibit the estimates obtained from generated values of Y_t (sample size = 200); the first column in the table shows starting values (namely consistent estimates) of α , σ_ε^2 and σ_η^2 , and the second column, the third, and so forth, show each step of the iteration.

TABLE 1(*)

Step of iteration	1st	2nd	3rd	4th	5th	6th	7th
α	0.705	0.703	0.701	0.700	0.699	0.697	0.696
σ_η^2	0.546	0.568	0.573	0.576	0.579	0.582	0.584
σ_ε^2	1.286	1.253	1.250	1.247	1.244	1.242	1.243

The estimation from 300 Y's yielded this:

TABLE 2(*)

Step of iteration	1st	2nd	3rd	4th	5th	6th	7th
α	0.517	0.523	0.526	0.529	0.532	0.533	0.535
σ_η^2	1.249	1.177	1.165	1.155	1.147	1.140	1.135
σ_ε^2	0.708	0.787	0.798	0.807	0.813	0.818	0.823

Obviously the case of sample size 300 gives better estimation than 200; but in both cases it can be observed that the convergence is very slow.

(*) In Tables 1 and 2, the values in the 1st column denote consistent estimates of α , σ_ε^2 and σ_η^2 which are starting values of iteration.

The following Table 3 displays the covariance matrices $\frac{4\pi}{N} W_0^{-1}$ of the estimates of α , σ_η^2 and σ_ε^2 for sample sizes N = 200 and 300 evaluated by means of the asymptotic covariance matrix given in Theorem 1, where the element corresponding to the column α and the raw σ_η^2 denotes, for example, the covariance of the estimates of α and σ_η^2 .

TABLE 3

N=200	α	σ_η^2	σ_ε^2	N=300	α	σ_η^2	σ_ε^2
α	0.050	-0.325	0.433		0.033	-0.216	0.289
σ_η^2	-0.325	2.178	-2.858		-0.216	1.145	-1.905
σ_ε^2	0.433	-2.858	3.811		0.289	-1.905	2.541

APPENDIX 1

THE NEWTON-RAPHSON METHOD

Let $L_N(\theta)$ be a likelihood function of $\theta = \{\theta_1, \theta_2, \ldots, \theta_q\}'$ given observations $X_1, X_2, \ldots, X_N$ and let $N_\delta(\theta^0) = \{\theta : \|\theta - \theta^0\| < \delta\}$ be a certain neighborhood of θ^0, the true value of θ. Now assume the following.

B-1) $\log L_N(\theta)$ is third-order differentiable with respect to θ_i, $i = 1, 2, \ldots, q$, for $\theta \in N(\theta^0)$.

$\ell_{ij} = \operatorname*{plim}_{N\to\infty} \partial^2 \log L_N(\theta^0)/N\partial\theta_i\partial\theta_j$ exists and the matrix

$\{\ell_{ij}\}$ $i, j = 1, 2, \ldots, p$ is nonsingular.

B-2) $\partial^3 \log L_N(\theta)/N\partial\theta_i\partial\theta_j\partial\theta_k$ is bounded in probability uniformly in $\theta \in N_\delta(\theta^0)$.

B-3) There exists a consistent estimate θ^1 [i.e., $\theta^1 \to \theta^0$ in probability as $N \to \infty$] such that $\sqrt{N}(\theta^1 - \theta^0)$ has a limiting distribution with a finite covariance matrix.

B-4) Let $\hat{\theta}$ be a solution of the likelihood equations which is consistent; then $\sqrt{N}(\theta - \theta^0)$ is also assumed to have a finite asymptotic covariance matrix.

Let $\Gamma_N(\theta)$ be a p by p matrix whose (i,j) element is

$\frac{\partial^2 \log L_N(\theta)}{\partial\theta_i \partial\theta_j}$ and $\gamma_N(\theta)$ be a p-vector whose i-th element is $\frac{\partial \log L_N(\theta)}{\partial\theta_i}$.

Let

$$\theta^2 = \theta^1 - \Gamma_N(\theta^1)^{-1} \gamma_N(\theta^1) \tag{15}$$

and

$$\theta^3 = \theta^2 - \Gamma_N(\theta^2)^{-1} \gamma_N(\theta^2) . \tag{16}$$

<u>Theorem 3</u>:

If B-1) through B-4) hold, $\sqrt{N}(\theta^2-\hat{\theta})$ tends to 0 in probability; in other words, $\sqrt{N}(\theta^2-\theta^0)$ has the same limiting distribution as the maximum-likelihood estimate. Furthermore, under the same conditions, $N(\theta^3-\hat{\theta})$ tends to 0 in probability.

<u>Proof</u>:

By the Taylor expansion of $\partial \log L_N(\hat{\theta})/\partial\theta_i = 0$ around θ^1,

$$\frac{\partial \log L_N(\theta^1)}{\partial\theta_i} + \sum_j(\hat{\theta}_j - \theta_j^1) \frac{\partial^2 \log L_N(\theta^1)}{\partial\theta_i\partial\theta_j} \tag{17}$$

$$+ \sum_j \sum_k(\hat{\theta}_j - \theta_j^1)(\hat{\theta}_k - \theta_k^1) \frac{\partial^3 \log L_N(\theta^*)}{\partial\theta_i\partial\theta_j\partial\theta_k} = 0 , \quad i = 1, 2, \dots, q,$$

where the θ^* is such that $\theta_i^1 \gtrless \theta_i^* \gtrless \hat{\theta}_i$ for $i = 1, 2, \ldots, q$. In (17) above,

$$\textstyle\sum_j (\hat{\theta}_j - \theta_j^1) \frac{\partial^2 \log L_N(\theta^1)}{\partial\theta_i \partial\theta_j}$$

(18)
$$= \textstyle\sum_j (\hat{\theta}_j - \theta_j^2) \frac{{}^2 \log L_N(\theta^1)}{\partial\theta_i \partial\theta_j} + \sum_j (\theta_j^2 - \theta_j^1) \frac{\partial^2 \log L_N(\theta^1)}{\partial\theta_i \partial\theta_j}$$

$$= {}_j (\hat{\theta}_j - \theta_j^2) \frac{\partial^2 \log L_N(\theta^1)}{\partial\theta_i \partial\theta_j} - \frac{\partial \log L_N(\theta^1)}{\partial\theta_i} ,$$

by (15). From (17 and (18),

$$\textstyle\sum_j \sqrt{N}(\hat{\theta}_j - \theta_j^2) \frac{\partial^2 \log L_N(\theta^1)}{N\partial\theta_i \partial\theta_j}$$

(19)
$$= -\textstyle\sum_j \sum_k \sqrt{N}(\hat{\theta}_j - \theta_j^1) \frac{\partial^3 \log L_N(\theta^*)}{N\partial\theta_i \partial\theta_j \partial\theta_k} \hat{\theta}_k - \theta_k^1) .$$

Writing the term on the right-hand side above as

$N(\hat{\theta}_j - \theta_j^1) \frac{\partial^3 \log L_N(\theta^*)}{N^{1+\epsilon}\partial\theta_i \partial\theta_j \partial\theta_k} N^{\epsilon}(\hat{\theta}_k - \theta_k^1)$, we see that, for $0 < \epsilon < \frac{1}{2}$,

both $\frac{\partial^3 \log L_N(\theta^*)}{N^{1+\epsilon}\partial\theta_i \partial\theta_j \partial\theta_k}$ and $N^{\epsilon}(\hat{\theta}_k - \theta_k^1)$ converge to 0 in probability

and $\sqrt{N}(\hat{\theta}_j - \theta_j^1)$ is asymptotically of finite variance by assumption.

Thus the whole quantity on the right of (19) converges to 0 in probability. It is easy to see that $\partial^2 \log L_N(\theta^1)/N\partial\theta_i\partial\theta_j$ converges to ℓ_{ij}. By assumption, the matrix (ℓ_{ij}) is nonsingular so that the $\sqrt{N}(\hat{\theta}_j - \theta_j^2)$ tends to 0 in probability. In order to prove the second assertion of the theorem, note the following equation:

$$
\begin{aligned}
N(\theta^3 - \hat{\theta}) &= \{I - \Gamma_N(\theta^2)^{-1}\Gamma_N(\hat{\theta})\} N(\theta^2 - \hat{\theta}) \\
&\quad + \Gamma_N(\theta^2)^{-1}\left\{\sum_{k=1}^{p} \partial\Gamma_N(\theta^{**})/\partial\theta_k(\theta_k^2 - \hat{\theta}_k)\right\} N(\theta^2 - \hat{\theta})
\end{aligned} \tag{20}
$$

where θ^{**} is a vector such that $\hat{\theta}_j \lesseqgtr \theta_j^{**} \gtreqless \theta_j^2$, $j = 1, 2, \ldots, p$, $\partial\Gamma_N(\theta)/\partial\theta_k$ is a p by p matrix with $\partial^3 \log L_N(\theta)/\partial\theta_i\partial\tau_j\partial\tau_k$ as its (i,j) element and I is the p by p identity matrix. Then if $N(\theta^2 - \hat{\theta})$ is bounded in probability, the first term on the right-hand side of (20) converges to 0 in probability since $\Gamma_N(\theta^2)^{-1}\Gamma_N(\hat{\theta})$ converges to the identity matrix, and the second term tends to 0 since $\Gamma_N(\theta^2)^{-1}\partial\Gamma_N(\theta^{**})/\partial\theta_k$ is asymptotically bounded. The fact that $N(\theta^2 - \hat{\theta})$ is bounded in probability is evident in view of the following equation:

$$
\begin{aligned}
N(\theta^2 - \hat{\theta}) &= \sqrt{N}\{I - \Gamma_N(\theta^1)^{-1}\,\Gamma_N(\hat{\theta})\}\,\sqrt{N}(\theta^1 - \theta) \\
&\quad - \Gamma_N^{-1}(\theta^1)\{\textstyle\sum_k \dfrac{\partial\Gamma_N(\tilde{\theta})}{\partial\theta_k}\sqrt{N}(\theta_k^1 - \hat{\theta}_k)\}\,\sqrt{N}(\theta^1 - \hat{\theta})
\end{aligned} \tag{21}
$$

where $\theta_j^1 \lesseqgtr \theta_j \lesseqgtr \theta_j$, $j = 1, 2, \ldots, p$, since

$$\sqrt{N}\{I - \Gamma_N(\theta^1)^{-1}\,\Gamma_N(\hat{\theta})\} = \sum_j \frac{\partial\Gamma_N(\tilde{\tilde{\theta}})}{N\partial\theta_j}\,\sqrt{N}(\hat{\theta}_j - \theta_j^1)$$

for a vector $\tilde{\tilde{\theta}}$ such that $\theta_j^1 \lessgtr \theta_j \gtrless \hat{\theta}_j$ and $\partial\Gamma_N(\tilde{\tilde{\theta}})/N\partial\theta_j$ is bounded in probability. $\square$

APPENDIX 2

THE ASYMPTOTIC PROPERTIES OF THE LEAST-SQUARES ESTIMATE OF A LINEAR PROCESS PLUS WHITE NOISE

Let $\{X_t : t = \ldots, -1, 0, 1, \ldots\}$ be a stationary process representable as a linear process $X_t = \sum_{i=0}^{\infty} \mu_i(\theta)\varepsilon_{t-i}$, where the ε_t are independent random variables such that $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2(\theta)$ and $E(\varepsilon_t^4) = C < \infty$. The μ_i and σ_ε are functions solely of $\theta = (\theta_1, \theta_2, \ldots, \theta_q)'$. Suppose that the process $\{X_t\}$ has a spectral density $f(\omega|\theta)$ with respect to the Lebesque measure. Then define a process $\{Y_t\}$ by $Y_t = X_t + \eta_t$ where $\{X_t\}$ is defined as above, the η_t are i.i.d. random variables with mean 0, variance $Var(\eta_t) = \sigma_\eta^2(\theta)$ and $E(\eta_t^4) < \infty$, and $\{\varepsilon_t\}$ is independent of $\{\eta_t\}$. Set $g(\omega|\theta) = f(\omega|\theta) + \frac{1}{2\pi}\sigma_\eta^2(\theta)$. Now assume the following:

C-1) θ^0, the true value of θ, is in Θ, a compact subset of R^q,

C-2) $g(\omega|\theta^1)$ cannot be equal to $g(\omega|\theta^2)$ i.e., for $\theta^1 \neq \theta^2$,

C-3) $h(\omega|\theta) = 1/g(\omega|\theta)$, and $h^{(i)}(\omega|\theta) = \partial h(\omega|\theta)/\partial\theta_i$, $i = 1, 2, \ldots, q$, are continuous in (ω,θ) for $|\omega| \leq \pi$, $\theta \in \Theta$, and W_0, the q by q matrix with the representative term

$$\int_{-\pi}^{\pi} \frac{h^{(i)}(\omega|\theta^0)}{h(\omega|\theta^0)} \frac{h^{(j)}(\omega|\theta^0)}{h(\omega|\theta^0)} d\omega$$

is nonsingular,

C-4) $h^{(i,j)}(\omega|\theta) = \partial^2 h/\partial\theta_i\partial\theta_j$ and $h^{(i,j,k)}(\omega|\theta) = \partial^3 h/\partial\theta_i\partial\theta_j\partial\theta_k$

exist and are continuous in (ω,θ) for $|\omega| \leq \pi$, and

$N_{\delta_1}(\theta^0)$, a neighborhood of θ^0 ; namely $N_{\delta_1}(\theta^0)$

$= \{\theta: ||\theta - \theta^0|| < \delta_1\}$,

C-5) $\sum_{i=0}^{\infty} i|\mu_i(\theta^0)| < \infty$.

Set $U_N(\theta) = \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\left|\sum_{t=1}^{N} Y_t e^{i\omega t}\right|^2 \Big/ g(\omega|\theta)d\omega$, and define $S_N(\theta)$

$$S_N(\theta) = -\frac{N}{4\pi}\int_{-\pi}^{\pi}\log f(\omega|\theta)d\omega - \frac{1}{2}U_N(\theta) .$$

Let $\hat{\theta}$ be a value of θ which maximize $S_N(\theta)$

Theorem 5:

Assume the conditions C-1 through C-5. Then $\hat{\theta}$, the approximate maximum-likelihood estimate of θ , is consistent, and $\sqrt{N}(\hat{\theta} - \theta^0)$ has the limiting distribution $N(0, 4\pi W_0^{-1} + K_4(\varepsilon) W_0^{-1}U_0W_0^{-1}) + K_4(\eta)W_0^{-1}V_0W_0^{-1})$, where W_0, U_0, V_0 are q by q matrices having i,j-th element

$$\int_{-\pi}^{\pi}\frac{h^{(i)}(\omega|\theta^0)}{h(\omega|\theta^0)}\,\frac{h^{(j)}(\omega|\theta^0)}{h(\omega|\theta^0)}d\omega ,$$

$$\int_{-\pi}^{\pi} h^{(i)}(\omega|\theta^0)f(\omega|\theta^0)d\omega \Big/ \int_{-\pi}^{\pi} h^{(j)}(\omega|\theta^0)f(\omega|\theta^0)d\omega ,$$

$$\frac{\sigma_\eta(\theta^0)}{2\pi}\int_{-\pi}^{\pi} h^{(i)}(\omega|\theta^0)d\omega \quad \frac{\sigma_\eta(\theta^0)}{2\pi}\int_{-\pi}^{\pi} h^{(j)}(\omega|\theta^0)d\omega$$

respectively. If the ε_t and η_t are Gaussian, the asymptotic distribution is $N(0, 4\pi W_0^{-1})$.

This theorem is derived by applying several modifications to Walker's results [1964]. For the arguments, the next lemma is important. The result is a straightforward extension of the Grenander-Rosenblatt theorem [1957, p. 137] and the proof is omitted.

Lemma 1:

Let $W_j(\omega)$, $j = 1, 2, \ldots, p$, be any bounded even functions of ω with at most a finite number of discontinuities; let $K_4(\varepsilon)$ and $K_4(\eta)$ be fourth cumulants of ε_t and η_t respectively. Then,

$$\lim_{N\to\infty} N \operatorname{cov}\left\{\int_{-\pi}^{\pi} I_N(\omega)W_j(\omega)d\omega, \int_{-\pi}^{\pi} I_N(\omega)W_k(\omega)d\omega\right\}$$

$$= 16\pi^2\left[4\pi\int_{-\pi}^{\pi} g(\omega)^2 W_j(\omega)W_k(\omega)d\omega + K_4(\varepsilon)\left\{\int_{-\pi}^{\pi} f(\omega)W_j(\omega)d\omega\right\}\left\{\int_{-\pi}^{\pi} f(\omega)W_k(\omega)d\omega\right\}\right.$$
$$\left. + K_4(\eta)\left\{\frac{\sigma_\eta^2}{2\pi}\int_{-\pi}^{\pi} W_j(\omega)d\omega\right\}\left\{\int_{-\pi}^{\pi} W_k(\omega)d\omega\right\}\right],$$

where $I_N(\omega)$ is the periodgram of the Y_t, namely $I_n(\omega) = \frac{1}{2\pi}\left|\sum_{\nu=1}^{N} Y_\nu e^{-i\nu\omega}\right|^2$.

A. The Consistency of $\hat{\theta}$.

Lemma 2 [Walker (1964) p. 368]:

There exists a function $H_{\delta,N}(\theta)$ of θ and $Y_1, \ldots, Y_n$ such that

$$|N^{-1}[U_N(\theta^2) - U_N(\theta^1)]| < H_{\delta,N}(\theta) \quad \text{for all} \quad \theta_1, \theta_2 \in \Theta(||\theta^2 - \theta^1|| < \delta) ,$$

$$\lim_{\delta \to 0} E(H_{\delta,N}) = 0 \quad \text{uniformly in } N ,$$

$$\lim_{N \to \infty} \mathrm{Var}\,(H_{\delta,N}) = 0 \quad \text{for each } \delta$$

Lemma 3:

Let θ^0 be the true value of θ , and θ^* be any other point in Θ , then,

$$\lim_{N \to \infty} P_r \left\{ \frac{1}{N} [S_N(\theta^0) - S_N(\theta^*)] > K'(\theta^0,\theta^*) \right\} = 1$$

for some positive $K'(\theta^0,\theta^*)$.

Proof:

$$\lim_{N \to \infty} E \left\{ \frac{1}{N} [S_N(\theta^0) - S_N(\theta^*)]\right\}$$

$$= -\frac{1}{2} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left\{ \frac{g(\omega|\theta^0}{g(\omega|\theta^*)} \right\} d\omega + \lim_{N \to \infty} E \left\{ \frac{U_N(\theta^0) - U_N(\theta^*)}{N} \right\} \right] ,$$

where the second term is equal to

$$1 - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g(\omega|\theta^0)}{g(\omega|\theta^*)} \, d\omega .$$

Thus

$$\lim_{N\to\infty} E \{\frac{1}{N} [S_N(\theta^0) - S_N(\theta^*)]\}$$

$$= -\frac{1}{2} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left\{\frac{g(\omega|\theta^0)}{g(\omega|\theta^*)} e\right\} d\omega - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left\{\exp\left[\frac{g(\omega|\theta^0)}{g(\omega|\theta^*)}\right]\right\} d\omega\right] .$$

But note that, for any $x \in R$, $xe \leq e^x$ with equality at $x = 1$. Hence

$$\log \left\{\frac{g(\omega|\theta^0)}{g(\omega|\theta^*)}\right\} e \leq \log \left\{\exp \left[\frac{g(\omega|\theta^0)}{g(\omega|\theta^*)}\right]\right\}$$

where, by condition 2, the equality does not hold for almost all ω. Thus

$$(22) \quad \int_{-\pi}^{\pi} \log \left\{\frac{g(\omega|\theta^0)}{g(\omega|\theta^*)} e\right\} d\omega - \int_{-\pi}^{\pi} \log \left\{\exp \frac{g(\omega|\theta^0)}{g(\omega|\theta^*)}\right\} d\omega < 0 .$$

On the other hand,

$$\mathrm{Var} \{ \frac{1}{N} [S_N(\theta^0) - S_N(\theta^*)]\} = \mathrm{Var} \{ \frac{1}{N} [U_N(\theta^0) - U_N(\theta^*)]\} ,$$

where the right-side term converges to 0, as $N \to \infty$, in view of Lemma 2. □

Define $H_{\delta,N}(\theta^1)$ as in Lemma 2 and let

$$J_\delta(\theta^1) = \max_{\{\theta:||\theta - \theta^1|| < \delta\}} \frac{1}{4\pi} \int_{-\pi}^{\pi} |\log g(\omega|\theta^1) - \log g(\omega|\theta) \, d\omega .$$

Now put $H^*_{\delta,N}(\theta^1) = H_{\delta,N}(\theta^1) + J_\delta(\theta^1)$.

Lemma 4:

Let $|\theta^2 - \theta^1| < \delta$, then

$$\left|\frac{1}{N}\left[S_N(\theta^2) - S_N(\theta^1)\right]\right| < H^*_{\delta,N}(\theta^1) \ , \quad \lim_{\delta\to 0} E(H^*_{\delta,N}) = 0$$

uniformly in N , and $\lim_{N\to\infty} \mathrm{Var}\,(H^*_{\delta,N}) = 0$ for any δ .

Proof:

$$\left|\frac{1}{N}\left[S_N(\theta^2) - S_N(\theta^1)\right]\right| \leq \frac{1}{2}\left|\frac{1}{N}\left[U_N(\theta^2) - U_N(\theta^1)\right]\right| + \frac{1}{4\pi}\int_{-\pi}^{\pi} |\log g(\omega|\theta^2)$$

$$- \log g(\omega|\theta^1)|d\omega \leq \frac{1}{2} H_{\delta,N}(\theta^1) + J_\delta(\theta^1) \ .$$

In view of Lemma 2, it suffices to prove that $\lim_{\delta\to 0} J_\delta(\theta^1) = 0$.

By the mean-value theorem,

$$\int_{-\pi}^{\pi} |\log g(\omega|\theta^1) - \log g(\omega|\theta^2)|d\omega$$

$$< \delta\int_{-\pi}^{\pi} \left|\sum_{k=1}^{q} \frac{\partial g(\omega|\lambda\theta^1+(1-\lambda)\theta^2)}{\partial\theta_k} h(\omega|\lambda\theta^1+(1-\lambda)\theta^2)\right| d\omega$$

where $\frac{\partial g(\omega|\lambda\theta^1+(1-\lambda)\theta^2)}{\partial\theta_k}$ and $h(\omega|\lambda\theta^1+(1-\lambda)\theta^2)$ are bounded functions on

$\{\theta^2 : |\theta^2 - \theta^1| < \delta\}$, by conditions 3) and 6). Thus as $\delta \to 0$

$$\max_{\{\theta\ |\theta - \theta^1| < \delta\}} \int_{-\pi}^{\pi} |\log g(\omega|\theta^1) - \log g(\omega|\theta)| d\omega \to 0 \ . \ \square$$

Now the consistency of $\hat{\theta}$ follows from Lemma 3 and Lemma 4, by almost the same steps given by Walker (1964).

B. Asymptotic Distribution of $\hat{\theta}$.

It holds that

$$\text{(23)} \qquad \sqrt{N}(\hat{\theta} - \theta^0) = - \left(\frac{\partial^2 S_N(\theta^*)}{N\partial\theta\partial\theta'} \right)^{-1} \frac{\partial S_N(\theta^0)}{\sqrt{N}\partial\theta}$$

where $\theta^* = \lambda\hat{\theta} + (1-\lambda)\theta^0$ for some λ , $0 < \lambda < 1$, $\partial^2 S_N(.)/N\partial\theta\partial\theta'$ denotes the q by q matrix with elements $\partial^2 S_N(.)/N\partial\theta_i\partial\theta_j$, $i, j + 1, 2, \ldots, q$; $\partial S_N(.)/ \sqrt{n}\partial\theta$ is the q-vector with elements $\partial S_N(.)/ \sqrt{n}\partial\theta_j$, $j = 1, 2, \ldots, q$ ($\partial^2 h(\omega|\theta)/\partial\theta\partial\theta'$ below is defined in the same way). By condition 4),

$$\text{(24)} \qquad \frac{\partial^2 S_N(\theta^*)}{N\partial\theta\partial\theta'} = - \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial^2}{\partial\theta\partial\theta'} \log h(\omega|\theta^*) d\omega$$

$$- \frac{1}{2(2\pi)^2 N} \int_{-\pi}^{\pi} \frac{\partial^2 h(\omega|\theta^*)}{\partial\theta\partial\theta'} |\Sigma_n X_n e^{in}|^2 d\omega \ .$$

Then, as $N \to \infty$, the right side of (24) converges to

$$-\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial^2}{\partial\theta\partial\theta'}\log h(\omega|\theta^0)d\omega - \frac{1}{2(2\pi)^2N}\int_{-\pi}^{\pi}\frac{\partial^2 h(\omega|\theta^0)}{\partial\theta\partial\theta'}|\Sigma_n X_n e^{in\omega}|^2 d\omega$$

in probability. As $N \to \infty$,

$$\frac{1}{2(2\pi)^2N}\int_{-\pi}^{\pi}\frac{\partial^2 h(\omega|\theta^0)}{\partial\theta_i\partial\theta_j}|\Sigma_n X_n e^{in\omega}|^2 d\omega \to \frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{h^{(i,j)}(\omega|\theta^0)}{} d\omega .$$

Also

$$\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log h(\omega|\theta^0)d\omega = \frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{h^{(i,j)}}{h} - \frac{h^{(i)}h^{(j)}}{h^2}\,d\omega .$$

Therefore the following result holds.

<u>Lemma 5</u>:

$$\frac{\partial^2 S_n(\theta^*)}{N\partial\theta_i\partial\theta_j} \to -\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{h^{(i)}h^{(j)}}{h^2}\,d\omega = -\frac{1}{4\pi}W_0 \quad \text{in probability.}$$

The only problem which remains is to find the asymptotic distribution of $\frac{1}{\sqrt{N}}\frac{\partial S_n(\theta^0)}{\partial\theta}$.

<u>Lemma 6</u>:

$$\lim_{N\to\infty} E\left[\frac{1}{\sqrt{N}}\frac{\partial S_N(\theta^0)}{\partial\theta_i}\right] = 0 \ , \quad i = 1, 2, \ldots, q \ .$$

Proof:

$$E\left(\frac{1}{\sqrt{N}}\frac{\partial S_N(\theta^0)}{\partial\theta_i}\right) = \frac{\sqrt{N}}{4\pi}\int_{-\pi}^{\pi}\frac{h^{(i)}}{h}\,d\omega - \frac{\sqrt{N}}{4\pi}\int_{-\pi}^{\pi}E\left|\frac{\Sigma_n X_n e^{i\omega n}}{\sqrt{2\pi}}\right| h^{(i)}d\omega$$

$$= \frac{\sqrt{N}}{2\pi}\int_{-\pi}^{\pi}h^{(i)}\int_{-\pi}^{\pi}\frac{\sin^2\frac{N}{2}(\lambda-\omega)}{2\pi N\sin^2\frac{\lambda-\omega}{2}}[f(\lambda)-f(\omega)]\,d\lambda d\omega$$

since

$$E\left|\frac{\Sigma_n X_n e^{i\lambda n}}{\sqrt{2\pi N}}\right|^2 = \frac{1}{2N}\int_{-\pi}^{\pi}\frac{\sin^2\frac{1}{N}(\lambda-\omega)}{\sin^2\frac{\lambda-\omega}{2}}g(\omega|\theta^0)\,d\omega\ .$$

According to Grenander and Rosenblatt (1957, p. 130),

$$\int_{-\pi}^{\pi}\frac{\sin^2\frac{N}{2}(\lambda-\omega)}{2\pi N\sin^2\frac{N}{2}}(g(\lambda)-g(\omega))\,d\omega = 0\left(\frac{\log N}{N}\right),$$

when $g(x)-g(y) = 0(|x-y|)$, and $f(\omega|\theta^0)$ has a bounded derivative (Walker (1964, p. 374)), so that

$$E\left(\frac{1}{\sqrt{N}}\frac{\partial S_N(\theta^0)}{\partial\theta}\right) = 0\left(\frac{\log N}{N}\right). \quad \square$$

Lemma 7:

$\frac{1}{\sqrt{N}}\frac{\partial U_N(\theta^0)}{\partial\theta}$ has the limiting distribution

$$N(0, \tfrac{1}{2}W_0 + \frac{1}{(2\pi)^2}\{K_4(\varepsilon)\,U_0 + K_4(\eta)\,V_0\})\ .$$

Proof:

The asymptotic normality follows from a similar argument to Walker (1964) p. 375. The asymptotic covariance is evaluated by setting $W_j(\omega) = h^{(j)}(\omega)$ in Lemma 1. $\square$

Now asymptotically, it holds that

$$\frac{1}{N}\frac{S_N(\theta^0)}{\partial\theta} \approx \frac{S_N(\theta^0)}{\sqrt{N}\partial\theta} - E\left(\frac{S_N(\theta^0)}{\sqrt{N}\partial\theta}\right) = \frac{1}{2}\frac{1}{\sqrt{N}}\left\{\frac{\partial U_N}{\partial\theta} - E\left(\frac{\partial U_N}{\partial\theta}\right)\right\} .$$

But in view of Lemma 7, the last term above is distributed as $N(0, \frac{1}{\pi} W_0 + \frac{1}{(2\pi)^2}[K_4(\varepsilon)\, U_0 + K_4\, (\eta)\, V_0])$. Accordingly in view of (23) and Lemma 5, $\sqrt{N}(\hat{\theta} - \theta^0)$ is asymptotically distributed as $N(0, 4\pi\, W_0^{-1} + K_4(\varepsilon)\, W_0^{-1}\, U_0\, W_0^{-1} + K_4(\eta)\, W_0^{-1}\, V_0\, W_0^{-1})$.

REFERENCES

Akheizer, N. I. (1956), Theory of Approximation, Frederick Ungar Publishing Co., New York.

Grenander, U. and Rosenblatt, M. (1957), Statistical Analysis of Stationary Time Series, John Wiley & Sons, Inc., New York.

Grether, D. M. and Nerlove, M. (1970), "Some properties of 'optimal' seasonal adjustment," Econometrica, 38, 682-703.

Hannan, E. J. (1970), Multiple Time Series, John Wiley & Sons, Inc., New York.

Hannan, E. J. (1971), "Non-linear time series regression, a discussion paper," Australian National University.

Hoffman, K. (1962), Banach Spaces of Analytic Functions, Prentice-Hall, Inc., New Jersey.

Hosoya, Y. (1977), "High-order efficiency in the estimation of linear processes," forthcoming in Ann. Stat.

Pagano, M. (1974), "Estimation of models of autoregressive signal plus white noise," Ann. Stat., 2, 99-108.

Parzen, E. (1967), "Time series analysis for models of signal plus white noise," Spectral Analysis of Time Series, 233-258, John Wiley & Sons Inc., New York.

Walker, A. M. (1964), "Asymptotic properties of least-square estimates of parameters of the spectrum of a stationary nondeterministic time series," Journal of the Australian Math. Soc., 4, 363-384.

Walker, A. M. (1969), "On the estimation of a harmonic component in a time series with stationary residuals," Proc. Int. Stat., 43, 374-376.

Whalen, Anthony (1971), Detections of Signals in Noise, Academic Press, New York and London.

Whittle, P. (1952), "Some result in time series analysis," Skandinavisk Aktuarietidskrift, Haft 1-2, 48-60.

Whittle, P. (1962), "Gaussian estimation in stationary time series," Bul. Inst. Int. de Statest., tome 39-2, Paris.

Whittle, P. (1963), Prediction and Regulation by Linear Least-Squares Methods, Van Nostrand, N. J.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations," Steffen L. Lauritzen, May, 1974.

18. "On the Interrelationships among Sufficiency, Total Sufficiency and Some Related Concepts," Steffen L. Lauritzen, June, 1974.

19. "Statistical Inference for Multiply Truncated Power Series Distributions," T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models in the Time and Frequency Domains," T. W. Anderson, June 1975.

21. "Asymptotic Properties of Some Estimators in Moving Average Models," Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting," Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear Discriminant Function," Chien-Pai Han, April 1976.

24. "Panels and Time Series Analysis: Markov Chains and Autoregressive Processes," T. W. Anderson, July 1976.

25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson, September 1976.

26. "The Recurrence Classification of Risk and Storage Processes," J. Michael Harrison and Sidney I. Resnick, September 1976.

27. "The Generalized Variance of a Stationary Autoregressive Process," T. W. Anderson and Raul P.Mentz, October 1976.

28. "Estimation of the Parameters of Finite Location and Scale Mixtures," Javad Behboodian, October 1976.

29. "Identification of Parameters by the Distribution of a Maximum Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.

30. "Discrimination Between Stationary Guassian Processes, Large Sample Results," Will Gersch, January 1977.

31. "Principal Components in the Nonnormal Case: The Test for Sphericity," Christine M. Waternaux, October 1977.

32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

TECHNICAL REPORTS (continued)

33. "Canonical Correlations with Respect to a Complex Structure," Steen A. Andersson, July 1978.

34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and I. Olkin, July 1978.

35. " Maximum likelihood Estimation for Vector Autoregressive Moving Average Models," T. W. Anderson, July 1978.

36. "Maximum likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.

37. "Efficient Estimation of a Model with an Autoregressive Signal with White Noise," Y. Hosoya, March 1979.